ARI Research Note 87-73

DTIC FILE COPY

AD-A188 636

A Comparison of the Army's Project A Cognitive and Psychomotor Tests to Analogous Air Force and Navy Tests

Mark Y. Czarnolewski

for

Selection and Classification Technical Area
Lawrence M. Hanser, Chief

MANPOWER AND PERSONNEL RESEARCH LABORATORY
Newell K. Eaton, Director

U. S. Army

Research Institute for the Behavioral and Social Sciences

December 1987

88 2 08 112

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Technical review by

Michael G. Rumsey

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POT, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

88 2 08 112

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

A188 636

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS: n/a
2a. SECURITY CLASSIFICATION AUTHORITY: n/a	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: n/a	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): ARI Research Note 87-73	5. MONITORING ORGANIZATION REPORT NUMBER(S): n/a
6a. NAME OF PERFORMING ORGANIZATION: U.S. Army Rsch Institute; 6b. OFFICE SYMBOL (If applicable): PERI-RS	7a. NAME OF MONITORING ORGANIZATION: U.S. Army Research Institute
6c. ADDRESS (City, State, and ZIP Code): 5001 Eisenhower Avenue, Alexandria, Va 22333-5600	7b. ADDRESS (City, State, and ZIP Code): 5001 Eisenhower Avenue, Alexandria, Va 22333-5600
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: n/a; 8b. OFFICE SYMBOL (If applicable): n/a	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: n/a
8c. ADDRESS (City, State, and ZIP Code): n/a	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
6.27.72.A	2Q162772A791	2.3.1	2.3.1.H.1

11. TITLE (Include Security Classification)

(U) A COMPARISON OF THE ARMY'S PROJECT A COGNITIVE AND PSYCHOMOTOR TESTS TO ANALOGOUS AIR FORCE AND NAVY TESTS

12. PERSONAL AUTHOR(S)

Mark Y. Czarnolewski

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Research Note	FROM 10/86 TO 10/87	December 1987	49

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Computer administered tests, Information processing, Navy visual-spatial reasoning tests, Paired-associate paradigm, Project A, Project BAT, Project LAMP

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report describes predictor batteries measuring cognitive abilities and psychomotor skills that are being developed by the research efforts represented by the Army's Project A, the Air Force's Learning Ability Measurement Program (LAMP) and Basic Attributes Test (BAT) Program, and the Navy's spatial abilities effort. Much of the enclosed information regarding the Air Force and Navy measures is based on R.E. Christal's reports on cognitive tests developed by TTCP Services (Christal, 1984, 1985). Comparisons made between Project A measures with those of the other services concentrate on: (a) conceptual comparisons (i.e., the construct representativeness and methodological paradigm on which the tests are based); and (b) operational comparisons, i.e., matching each Project A test to similarly constructed tests developed by the other two services).

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: LAWRENCE M. HANSER, Chief, SCTA	22b. TELEPHONE (Include Area Code): (202) 274-8275; 22c. OFFICE SYMBOL: PERI-RS

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

FOREWORD

The Army is currently engaged in a large-scale, long-term effort, known as Project A, to develop selection and classification measures that are directly linked to future Army performance. Potential selection and classification measures already developed cover four broad predictor domains-cognitive, psychomotor, temperament and interest. Army developmental efforts in the cognitive and psychomotor domains parallel to some extent recent developments in the U.S. Air Force and Navy. This report presents for the first time an analytic integration of the cognitive and psychomotor measures developed by the three services, building upon recent work by R. E. Christal. It provides a catalogue of measures available in each of the major services and compares measures on selected characteristics. The information provided should be particularly useful to potential test users, both in the Army and in the other services, in the initial identification and review of tests for particular measurement purposes, and to researchers interested in the test domains covered. However, this document should not be relied upon for exhaustive treatment of each of the tests covered; those interested in more detailed information should consult the relevant documentation provided by the developing agency.

Project A was authorized through a Letter, DCSOPS, "Army Research Project to Validate the Predictive Value of the Armed Services Vocational Aptitude Battery", effective 19 November 1980; and a Memorandum, Assistant Secretary of Defense (MRA&L), "Enlistment Standards", effective 11 September 1980.

In order to ensure that Project A research achieves its full scientific potential and will be maximally useful to the Army, a governance advisory group comprised of Army General Officers, Interservice Scientists, and experts in personnel measurement, selection and classification was established. Members of the latter component provide guidance on technical aspects of the research, while general officer and interservice components oversee the entire research effort, provide military judgment, provide periodic reviews of research progress, results and plans, and coordinate within their commands. Recently members of the General Officer's Advisory Group included MG W.G. O'Leksy (DMPM) (Chair), MG C.F. Briggs and MG T.J.P. Jones (FORSCOM, DCSPER), BG W.C. Knudson and MG J.B. Allen, Jr. (DCSOPS), BG F.M. Franks, Jr. (USAREUR, ADCSOPS), and MG J.H. Corns (TRADOC, DCS-T). The General Officers' Advisory Group has been briefed periodically, from 1983 to, most recently, May, 1987, on the development and data pertaining to the Project A cognitive and psychomotor predictor measures described in this report. Members of Project A's Scientific Advisory Group (SAG) guide the technical quality of the research. During the period covered by this report, they included Drs. Philip Bobko, Thomas Cook, Milton Hakel (Chair), Lloyd Humphreys, Robert Linn, Larry Johnson, Mary Tenopyr, and Jay Uhlaner. The SAG has been briefed biannually, from 1983 to, most recently, September 1987 on the Project A predictor measures described in this report.

FOREWORD (CONTINUED)

A comprehensive set of new selection/classification tests and job performance/training criteria have been developed and field tested, and the revised tests have been administered in a large-scale concurrent validation data collection effort. Results will be used to link enlistment standards to required job performance standards and to more accurately assign soldiers to Army jobs.

ACKNOWLEDGEMENTS

Acknowledgements are extended to Frances Grafton, Stephen Cormier, Elizabeth Smith and Martha Teplitzky for providing helpful and constructive feedback, and to Deborah Lucas, Betty Shelly and Joanne Proctor for typing and reformatting the layout of the document. Special thanks to Clessen Martin and Jane Arabian for their consistently constructive editorial comments and suggestions.

A COMPARISON OF THE ARMY'S PROJECT A COGNITIVE AND PSYCHOMOTOR TESTS TO ANALOGOUS AIR FORCE AND NAVY TESTS

EXECUTIVE SUMMARY

Requirement:

To compare recent tests developed by the Air Force and Navy with that portion of the Army's Project A predictor battery that measures cognitive and psychomotor aspects of human performance. The purpose was to identify areas of overlap and uniqueness between the Army tests and the tests of the other two services.

Procedure:

The author reviewed two tri-service reports (Christal, 1984, 1985) and reconstructed descriptions of the Air Force Human Resources Laboratory and Naval tests into tables describing the tests in terms of the following: a brief verbal description, experimental design, statistics representing subject performance, reliability and construct representativeness. Tables detailing each test's similarity to the Army's Project A predictor battery are included.

Findings:

A number of tests being developed by the other services appear to be similar to the Project A predictor battery, (e.g., Project A's Number Memory and Memory Scanning Tests). However, there appear to be tests unique to each service; of particular note is the use of the paired-associate paradigm for a number of tests developed by the Air Force. Tests that neither service appears to be stressing or developing in the programs covered by the present report include selective attention measures.

Utilization of Findings:

The findings suggest that the Project A predictor battery includes measures of common interest with the other services, but that there may be measures which one or the other of the services is developing which may be candidates for supplements to the Project A Battery. These include measures that focus on (a) handling information overload, (b) handling a continual stream of competing information, or (c) capturing learning rates of tasks that appear to be essential components of one or more real world tasks. A data base that defines tests in terms of the primary test characteristics identified by this report may allow for more systematic listing of tests for determining further candidate measures of potential use in Army selection and classification testing.

A COMPARISON OF THE ARMY'S PROJECT A COGNITIVE AND PSYCHOMOTOR TESTS TO ANALOGOUS AIR FORCE AND NAVY TESTS

CONTENTS

LIST OF TABLES

CONTENTS (CONTINUED)

INTRODUCTION

Mark Y. Czarnolewski and Michael G. Rumsey

Psychology has recently seen an explosion of measures which capture reliable individual differences in performance in the content domains of cognitive psychology, psychomotor performance and ability testing (Hunt & Pellegrino, 1985). These new measures have operationalized constructs that appear sensitive to a person's ability to use different types of information (e.g., verbal versus physical types of information) or to a person's ability to use search strategies when making a decision based on that information (e.g., exhaustive versus self-terminating decision-based searches).

Many of these new measures of individual differences represent more dynamic testing situations than the paper and pencil individual difference measures classically found in the literature. Microcomputers are increasingly becoming more adapted to mass testing of these dynamic situations, with the result that a variety of these measures can be used for either selectio and classification or diagnostic purposes.

The military services are currently developing and augmenting their test batteries for both predictive and diagnostic purposes. Christal (1984, 1985) lists tests being developed by the U.S. Army, Air Force, and Navy, a separate Tri-Services research effort that has input from the above three services, and military services of other countries, most notably Canada and England. Christal's reports provide lists of tests and appear to allow each of the services covered to present information of their developing tests in a format selected by that service.

Christal's reports, however, do not attempt to look at the interrelationships among the tests. The identification of tests is a worthwhile service, but greater understanding of the services' testing programs and the interrelationships among the described tests is needed. There is a need to identify test characteristics which allow one to compare and contrast the different tests.

The present report attempts to compare a number of the tests in Christal's reports in terms of the operational procedures that appear to be required to perform the tests and in terms of the constructs that the tests are designed to represent. The report also refers to methodologically sound and theoretically based concepts so that the reader can get a "handle" on the large number of diverse measures that are described in Christal's reports.

This report will reflect the investment into test development from the U.S. Army's perspective; specifically, the report will briefly describe the research efforts and specific tests being developed by the U.S. Army, Air Force and Navy. The report will then attempt to focus on the Army measures by comparing these measures to the measures of the other two services.

The Army research project which is developing the predictors reviewed in this report is called Project A (Eaton, Goer, Harris & Zook, 1985). Project A is designed to evaluate and update the selection and classification system of the Army. The project entails the development of both

criterion and predictor measures, with the intent that the developed predictor measures and the currently used predictor measures (the latter being the Armed Services Vocational Aptitude Battery or ASVAB) be validated against the criterion measures (Eaton et al., 1985). The intent of the present report is to describe a portion of the research effort of Project A and to compare the described portion of Project A to similar efforts of the Air Force and Navy. Specifically, of the four original predictor constructs of interest in Project A -i.e., Cognitive, Psychomotor, Temperament and Interest (Peterson, 1985)- the present paper will deal with just the first two.

This report is not intended to be a comprehensive literature review of cognitive measures or psychomotor skills, in general, or of all Air Force, Navy and Army tests designed to represent these domains of behavior. Rather, the intent is to restructure the extensive verbal descriptions of many of these tests that are described in Christal's reports into a tabular format that facilitates direct comparison. The tables contain brief descriptions that highlight relevant task and/or psychometric characteristics of each test.

The structure of the report reflects the intent of preliminary evaluation and across service comparisons. First, each service's test development program(s) is(are) described briefly, and each program's tests are described and critiqued. Afterwards, comparisons are made between Project A tests and the Air Force and Navy tests. These comparisons include discussions of the research activities of these three services and categorization of the research activities as "similar", "different", or "not dealt with". Concluding comments suggest that the perceived overlaps and distinctions among the services may be reflecting the common and unique research questions that each service seeks to answer.

It should be noted that this report labels both ability tests and information-processing tasks as cognitive measures. It should also be noted that a variety of approaches have been used to study such measures. One approach attempts to further the understanding of a construct by exploring the psychometric characteristics of a test that is developed to represent that construct. Using this psychometric approach, researchers have, for example, developed reliability techniques to evaluate a test's measurement consistency, and have developed factor analytic techniques to define particular constructs and to test their validity. A second approach uses experimental manipulations to examine human behaviors associated with particular constructs. From this approach have come methodologies to represent learning strategies, as well as search and other decision-making strategies.

The recent integration of the two previously disparate areas of psychometrics and experimental psychology is found in the study of individual differences in learning or information processing tasks. This integration is exemplified by Hunt's (1978) work, which factor analyzes a number of experimental tasks to construct a model of attention, or by Sternberg (1977), who employs an experimental manipulation of an inductive reasoning

task; a task which, in the past, would have been described and constructed in terms of psychometric techniques. Identification of individual differences may, thus, be achieved by either psychometric or experimental approaches for studying human behavior. The present paper attempts to capture the more integrative approach by reporting, where possible, information on a given task or test in terms of psychometric and experimentally-based descriptors.

The author of this report must, however, acknowledge an inconsistency in the coverage of the tests in this report. One inconsistency is due to the reference documents' (i.e., Christal, 1984, 1985) not employing the same descriptions for evaluating each test's properties (e.g., psychometric properties). The other inconsistency is due to the author's desire to discuss the paradigms on which some of the tests are based to highlight unique characteristics of the tests or to comment on the ways those tests are or are not being used. For example, there is a discussion of a subject's performance changes when taking tests that are given over repeated blocks of trials. This discussion is found in a section describing measures of automaticity - a paradigm in which extensive practice results in a qualitative change in the strategy a subject employs when performing certain tasks. The author briefly describes the theory involved in the automaticity paradigm and points out other possible experimental strategies than the ones currently being considered by the service using the paradigm (e.g., the Air Force). Such a discussion would not be relevant for all tests. Comparisons between measures developed by psychometric methodologies and measures developed by experimental methodologies are also found only in sections where such discussions are relevant, such as the Air Force learning curve parameter identification program. The research focus of that program suggests that one may define a given learning task in terms of an ability that is defined by a test from the psychometric tradition, or by a cognitive operation from the experimental tradition -- a point which, again, would not be relevant for all tests covered by this report.

In summary, this report is intended to do the following:

- Describe cognitive or psychomotor measures being employed by primary research efforts in the Army, Air Force and Navy. These include Air Force and Navy tests as of 1985 and Army tests as of 1986.

- Provide the reader with some background on the paradigms or theories on which some of the tests are based;

- Where the available background information permits, specify particular ways in which the tests might be used as predictors or diagnostic tools;

- Compare the Army's Project A cognitive and psychomotor measures to the Air Force and Navy tests in terms of each test's stimulus characteristics, construct representativeness, or paradigmatic foundation;

- Identify similarities and differences between Project A cognitive and

psychomotor measures and the Air Force and Navy cognitive and psychomotor measures; the discussions on differences include (a) use of different tests and (b) use of similar tests differently (e.g., more or less practice).

AIR FORCE PROGRAMS

The Air Force Human Resource Laboratory (AFHRL) research programs appear to stress identification of individual differences in a number of tasks representative of both the learning and information processing facets of the experimental literature. Two of the programs are the Basic Attributes Test (BAT) Program and the Learning Abilities Measurement Program (LAMP). A brief description of each program and its respective tests follows. The LAMP project is the more extensive of the two programs; consequently, most of this section will discuss LAMP.

Basic Attributes Test Program

The Basic Attributes Tests (BAT) were designed to predict aircraft pilot performance. Two behavioral domains represented include personality assessment and information-processing "capability" identification. The intent for the latter domain was to identify tasks that measure the abilities to quickly sort, prioritize and act on a continual stream of visual, auditory and tactile information. The personality measures of the BAT are included in the present review of cognitive and psychomotor measures because a number of the personality measures are of a dynamic nature and require computer administration. They were, therefore, considered relevant to this review.

The BAT tests are listed in Table 1. Numbers are used to identify a test's location in the table. Sections of the table are not filled in because relevant information was not provided by the Christal reports. This format is applied for all tables in this report.

The tests in Table 1 include, from the classical psychometric tradition, abilities adapted for computer testing, such as Perceptual Speed (1) and Field Dependence (10). Information processing paradigms from the 1970s include encoding speed (4), mental rotation (5) and item recognition (6). Channel capacity models from the 1960s are represented by Time-Sharing (3), which includes the ability to simultaneously process information load and tracking information, and Decision-Making Speed (8), which requires the ability to process an increasing amount of information load. A psychomotor task (16) is also included.

The BAT's label of Encoding Speed for one of its tests deserves comment. This task was developed by Posner (Posner, Boies, Eichelman & Taylor, 1969) and may be more accurately described as a SAME-DIFFERENT judgment task that distinguishes between nominal and physical identity (Nickerson, 1972; Posner, 1978). The term "encoding" is usually employed to represent one of the hypothesized information processing stages for this task (Farrell, 1985; Krueger, 1978, Nickerson, 1972).

Table 1

Air Force Basic Attributes Test

Task Name	Task Description	Design	Statistic of Interest
1. Perceptual Speed	Press buttons in same order as sequential presentation of digits.	Digit Size = 4	
2. Dot Estimation	Decide which of two boxes has more dots (Compulsive vs. Decisive Behavior).		
3. Time Sharing	Maintain a marker's position in a compensatory task.	Blocks = 3; 5 trials/Block Blocks (No secondary task vs. Secondary task requiring observing digits and pressing corresponding buttons vs. Same secondary task but with time constraint).	
4. Encoding Speed	Posner task requiring SAME-DIFFERENT judgments.	Type of Identity (Physical vs. Nominal vs. Categorical, e.g. consonant or vowel).	
5. Mental Rotation	Task requires SAME-DIFFERENT judgments.	Perspective (SAME vs. Mirror Image); Orientation (SAME vs. Rotated).	
6. Item Recognition	Sternberg Task - Decide whether a digit was present or absent in a previous array.	Memory Load (1 to 6 digits).	
7. Immediate/ Delayed Memory	Given a number in a sequence of numbers select the number that is one or two steps back (Modified Sternberg Task).	Memory Load (not specified). Backtrack from Specified Item (1 vs. 2 steps).	

Table 1 (Continued)

Air Force Basic Attributes Test

Task Name	Task Description	Design	Statistic of Interest
8. Decision-Making Speed	Choose correct response in a simple RT task with a 1:1 mapping between stimulus and response	Amount of Uncertainty (2 vs. 4 vs. 8 potential stimuli) x Type of Information (where and when signal will occur vs. where signal will occur vs. when signal will occur vs. neither where nor when signal will occur.	
9. Risk-Taking	Out of 10 boxes, pick as many boxes one-at-a-time without picking the "disaster" box.		
10. Embedded Figures	Decide which of 2 complex geometric figures contains a specified simple figure.		RT, Errors
11. Self-Crediting Word Knowledge	Project own performance in a vocabulary task of increasing difficulty.	Blocks = 3, 10 questions/ Block. Increase Difficulty (Block 1 vs. Block 2 vs. Block 3).	
12. Activities Interest Inventory	Decide interest between activities differing in danger to self.	Pairs of activities = 81.	
13. Automated Aircraft Personality Profiler	Decide preferred attitude and interest.	Number of paired choices = 200.	
14. Biographical Data			Identity, Age Gender, etc. Personal History. Attitudes Toward flying
15. Psychomotor-a	Two-handed coordination		
16. Psychomotor-b	Coordinate stick and rudder		

Personality type measures include Dot Estimation (2) Risk Taking (9), Self Crediting Word Knowledge (11), Interest (12), and Personality (13) profiles.

Learning Abilities Measurement Program

The Learning Abilities Measurement Program (LAMP) is a long-term basic research program intended to develop a theory-based system of ability measurement. The program relies on microcomputers and other measurement devices to measure previously unevaluated abilities, response latencies and learning efficiencies under laboratory-controlled conditions.

Four programs are embedded within the Learning Abilities Measurement Program (LAMP). These are:

(a) tasks used to establish a baseline of information on Air Force enlisted personnel;
(b) learning tasks used to evaluate learning curve parameters in predicting subsequent learning proficiencies;
(c) tasks designed to measure individual differences in movement toward automaticity;
(d) tasks constructed for use in various in-house experiments (Christal, 1984).

Each of these programs will be discussed in turn.

Baseline Tasks

Table 2 contains a list of the information-processing and learning tasks. The table contains a short description of each task, the experimental design embedded within each task, the statistics employed to represent subject performance and the reliability of those statistics.

The tasks include simple reaction time (RT) (1) and binary classification tasks, such as choice RT (2), variations on the Posner Nominal-Physical Match Task (3,4), Sternberg Memory Scanning Task (7a, 7b), and Neisser's Search for Presence Task (21). There is also an emphasis on paired-associate tasks of various forms. These include the basic paradigm (10), retest format (13), retention despite interference (i.e., retroactive interference) (14), and updating of information (19). Another memory test is the missing digit task (17). Inductive reasoning tasks are represented (18, 20) as is a deductive reasoning task (16). Tasks evaluating the effect of sentence structure on comprehension are also represented (9, 28). Construct comparisons between these measures and those of Project A will be found in the section entitled "Army's Project A Predictor Battery".

Methodological Issues. An initial comparison will be made between the Air Force Tests and the Project A tests to highlight the test characteristics that were reviewed to both describe the tests and to use as standards for comparing the tests. The description of the Baseline test battery is

Table 2

Air Force Baseline Tests

Task Name	Task Description	Design	Data and Reliability: Statistic of Interest	Reliability Type	Amount
1. Simple RT		Blocks = 6; 20 trials/Block	Mean RT	Direct Computation	.99
		Blocks 1,3,5 right forefinger		Odd/Even Split-half,	
		Blocks 2,4,6 left forefinger	S.D. RT	Corrected for length.	.99
				Odd/Even Split-half, (cfl)	.99
2. Two Choice RT		Blocks = 6; 20 trials/Block	Mean RT	Method of Average:	
				Correlation, cfl	.96
				Direct Computation	.98
3. Posner Physical Match (PI)		Blocks = 2; 80 trials/Block.	Mean RT	Test-Retest, cfl	.97
4. Posner Name Match (NI)	Within each Block, P.I. = 43 trials, N.I. = 37 trials	Blocks = 2; 80 trials/Block. with each Block	Mean RT	Immediate Test-Retest	.96
			NI - PI	Reliability Difference Scores	.43
5. Single Word Categorization	Categorize words connotation	Blocks = 6; 20 trials/Block. within each Block, 1/2 trials positive and 1/2 trials negative connotation	Decision Time	Split-half, cfl	.93
				Direct Computation	.99
			Number Errors	Split-half, cfl	.75
			Categorization Time-Two Choice RT	Reliability of Difference scores	.89
6. Word Pairs	Pairs of Words Presented Simultaneously	Blocks - 2; 80 trials/Block. Type of connotation, Positive = SAME Negative - DIFF	Mean RT	Split-half, cfl	.99
7a. Sternberg	Memory Scanning Numbers	Memory Load (2,3,4,6) 160 Trials	Overall MeanRT	Direct computation	.98
			Slope Across Load	Split-half, cfl	.83
			Intercept of Slope	Split-half, cfl	.88
7b. Sternberg	Memory Scanning Words; examinees receive 4 consecutive trials at a given set size for 7a, 7b.	Memory Load (2,3,4,6)	Overall mean RT	Direct Computation	.99
		160 Trials	Slope Across Load	Split-half, cfl	.87
		Intercept of Slope		Split-half, cfl	.94
8. Numerical Operations	A separate set for each: addition, subtraction, miltiplication, division, mixed.	Sets - 5, 36 Trials/set	Add. Mean RT	Reliability	.95
			SD RT	Reliability	.53
			Subt. Mean RT	Reliability	.95
			SD RT	Reliability	.50
			Mult. Mean RT	Reliability	.91
			SD RT	Reliability	.59
			Divs. Mean RT	Reliability	.93
			SD RT	Reliability	.71
			Mixed Mean RT	Reliability	.97
			SD RT	Reliability	.75

Note. cfl = corrected for length

Table 2 (Continued)

Air Force Baseline Tests

Task Name	Task Description	Design	Data and Reliability: Statistic of Interest	Reliability Type	Amount
9. Sentence Verification	Determine truth or falsity of simple sentences of the following: X O The X is not follwed by the O.	Sets = 4, 32 sentences/set, Factorial Design: "To follow" vs. "To Precede" Active vs. Passive Positive vs. Negative Form O first vs. X first	Items Correct	Average Subtest Intercorrelation, cfl	.95
			Mean Time to Complete Form	Average Subtest Intercorrelation	.94
10. Paired Associates Learning	Learn a single digit number associated with a low associ-CVC trigram	10 associations; Criteria: a) get 10/10 correct in any 2 of 3 consecutive blocks or b) continue until 25th block	Number of trials required to learn individual word	Coefficient Alpha	.78
11. Moyer-Landauer	Subject decides the larger of two horizontally adjoining digits.	Sets = 2; 72 digit pairs/set	Mean RT	Test-Retest, cfl	.99
			Slope = func. (digit diff)	Test-Retest, cfl	.56
			Intercept of "diff" slope	Test-Retest, cfl	.97
			R-Square of "diff" slope	Test-Retest, cfl	.47
			Slope = func. (digit ratio)	Test-Retest, cfl	.54
			Intercept of "ratio" slope	Test-Retest, cfl	.96
			R-Square of "ratio" slope	Test-Retest, cfl	.49
12. Choice RT	Subjects primed for Number of stimuli, with 1:1 stimulus: response mapping. Captures handling of information transmission, per se.	Blocks = 5, 80 trials/Block. Stimulus Load (2,3,4,6)	*Slope RT as a function* of log-base-2 (N+1)	Reliability	.73
			Intercept of slope	Reliability	.70
			Correlation of slope	Reliability	.36
13. Relearning of Paired Associates	Re-administer Task 10 and compute proportinal reduction in learning trials	See Task 10.	Proportional reduction in learning trials	Coefficient Alpha	.46
			Trials required to learn	Coefficient Alpha	.77
14. Old-New Word Memory	Present one word at a time and subject decides if word was presented before.	Trials = 200 common nouns. No. of intervening words (0,1,2,3,4,5).	RT for lag and new words	Direct computation, the average of six reliabilities	.89
15. Random Production	Tap two keys, with respective fore-fingers. Each sequence of tappings are subjectively chosen.	Four sequences of 120 responses.	All of the following were computed by the avg. subset intercorrelation, cfl *phi-1=.66;* phi-2=.91; phi-3=.81; phi-4=.72; phi-5=.54		
16. Three-Term Series	Compare and make a choice based on information in two consecutive sentences. e.g., John not as tall as Pete, Pete not as tall as John. Who is shortest 1) Dick 2) John 3) Pete.	Sets = 3; 32 problems/set. Set 1: Dick, John, Pete; short-tall. Set 2: Dick, John, Pete; good-bad. Set 3: Red, Blue, Yellow; rough-smooth.	Response Time	Mean subtest intercorrelation, cfl	.92
			Number Correct	Mean subtest intercorrelation, cfl	.85
			Test Internal Consistency	Mean KR-21 for the three subtests	.85

Table 2 (Continued)

Air Force Baseline Tests

Task Name	Task Description	Design	Data and Reliability: Statistic of Interest	Reliability Type	Amount
17. Memory Span	Decide which one of previously (successively) presented nine digits is missing.	Blocks = 4; 27 Trials/Block	Computed memory span	Avg-subtest inter-correlation, cfl	.90
			Computed span over four Blocks	Direct computation	.98
			RT	Direct computation	.99
18. Patterns in Letter Series	Out of four groups of letters, detect and identify anomalous group.		RT	Direct computation	.85
			Internal Consistency	KR-21	.76
			Number Correct	Split-half, cfl	.83
19. Continuous Paired Associates	Subjects associate temperatures with each of four cities, where temperatures are frequently changing.	Blocks = 4, 50 Trials/block lag (0 Thru 8) - number of temperatures between presentation of a city's temp. and the question about the temp.	RT	Direct Computation	.85
			Proportion Correct	Split-half, cfl	.90
			Slope Prop. Cor. regressed on lag	Split-half, cfl	.73
			Intercept	Split-half, cfl	.94
20. Simon and Kotolvsky (Last letter of Pattern)	Choose the letter which continues a series of letters.	15 test questions.	Internal Consistency of right-wrong	KR-21	.58
21. Visual Scan (Neisser's Presence Task)	Search for a letter in a matrix of letters, with rows = 20 and columns = 7	Blocks = 4, 20 Trials/Block	Scan Rate	Direct Computation	.97
			Correct Identification	Average Subtest intercorrelation cfl	.89
22. Line Length Judgment	Pick the longer of two contiguous horizontal lines	Blocks = 4, 32 Trials/Block	d'	Test-Retest, cfl	.73
			beta	Test-Retest, cfl	.63
23. Sunday-Tuesday Task	Perform symbolic arithmetic with Monday = 1 and Sunday = 7. Answers are days of the week, not numbers	Sets = 2, 49 problems/set	Mean RT	Test-Retest, cfl	.94
			S.D. RT	Test-Retest, cfl	.82
			No. of items correct	KR-20	.96
24. Semantic Priming	Present two words and decide if both words are real.	Blocks = 3, 102 pairs/Block Both words (pairs=34). Both nonwords (pairs=34). One word real (pairs=68). Both words real and semantically related, e.g., doctor/nurse (pairs=102). Both words real and not semantically related (pairs=102)	RT difference btwn related and unrelated pairs	Splt-half, cfl	.87
				Reliability of difference scores	.76
25. Rotated Figures	Judge whether a rotated figure is the same as an upright figure on its left.	Trials = 200. Decision (Same - 120 Trials vs. Diff = 80 trials). Rotation (50,80,150 degrees).	RT slope across rotation	Split-half, cfl	.76

Table 2 (Continued)

Air Force Baseline Tests

Task Name	Task Description	Design	Statistic of Interest	Reliability Type	Amount
			Data and Reliability		
26. Dual Task	Continuously tap at a steady rate while performing secondary tasks of varying difficulty requiring searching for a v in a letter array.	Conditions = 8, 4 in ascending order of difficulty and the same 4 descending 1) baseline, 2) letter array with O's, 3) array with curved letters, 4) array with angular letters.	Michon perceptual motor load = (loaded tapping-unloaded tapping)/ unloaded tapping.	Avg. Subtest intercorrelation, cfl	.85
27. Ichakawa Short-term Memory for Visual Position	Respond if a pattern of 3 points formed by successively presenting 1 point in each of 3 (5 pt. x 5 pt.) matrices was prompted by a prior similar criterion matrix with 7 randomly selected matrix points.	Sets = 2, 30 Trials/set Presentation delay is 1, 5 and 9 sec. after off set of original matrix.	Number Correct Answers	Split-half, cfl	.89
			Slope of No. of items correctly answered regressed delay	Split-half, cfl	.24
			Intercept	Split-half, cfl	.77
28. Collins and Quillian	Answer true/false to a sentence whose subject and predicate differ in linguistic and cognitive complexity.	Sentences = 208 P sentences refer to subject properties, e.g., Wool is soft. S sentences refer to category membership, e.g., Christmas is a holiday. Level difference (0,1,2).	Slope of RT regressed on difference in levels	Odd/even split, cfl for S condition	.44
			Difference btwn Level 0 and Level 1 RT	P condition not significant	
				P condition relationship of difference	.10
				S condition relationship of difference	.58

more detailed than the previously reviewed BAT battery because the documents reviewed for this report (i.e., Christal, 1984, 1985) did not provide as much detail of test characteristics for the BAT battery as for the Baseline test battery. It was, thus, felt that this initial comparison, which is designed to sensitize the reader to the test character-istics that will be used to compare the Project A tests to the Air Force and Navy tests, would best be served when describing the Air Force Baseline tests.

Baseline tests contain more trials per task than found in Project A. For example, the Air Force's Sternberg task, using digits for stimuli (7a), contains 160 trials while the Sternberg task in the Project A battery contains 36 trials, with 1/3 of the stimulus trials containing digits.

As seen in Table 2, the Air Force's method to assess task reliability is <u>not</u> constant across tasks. Split-half reliabilities based on odd-even (parallel) items are computed for some tasks while other methods were used for other tasks. The "Direct Computation" method is not defined in Christal (1984).

Project A has employed test-retest and split-half reliability measures for its computer tasks (Toquam, Dunnette, Corpe, McHenry, Keyes, McGue, Houston, Russel & Hanson, 1985). Much concern has been directed toward the psychomotor tasks because these measures traditionally show low test-retest reliabilities.

The issue of adequate test-retest reliability may be reconceptualized as consistency in performance over many trials (Rogosa, Brandt & Zomowski, 1982; Tucker, 1966; Zeaman & Kaufman, 1955). Low reliabilities may be reflecting individual differences in learning rate for these tasks. One issue of possible interest is whether statistics reflecting subject learning may provide reliable and valid indicants of predictor performance.

<u>Tasks used to evaluate learning curve parameters</u>.

Table 3 contains learning tasks used to evaluate curve parameters in predicting subsequent learning proficiency. The table appears to contain job-relevant criterion tasks that are administered via a paired-associate format. For example, the Emergency Procedures Task requires that subjects learn critical (job-related) steps and those steps' relative order of performance. The task is structured so that one can repeatedly measure a subject's performance as he/she is learning the task.

A number of positive points may be made regarding the Air Force's research strategy of identifying tasks and their respective critical behaviors, and restructuring those tasks to fit an experimental paradigm. These, first, include the employment of sensitive dependent measures. For example, one can measure RT to the millisecond and distinguish between errors of commission and omission. Second, one may vary the tasks by employing a componential approach or introduce other independent variables. Third, one may impose various experimental frameworks onto the task, e.g.,

Table 3

Air Force Learning Tasks Used to Obtain Learning Curve Parameters

Task Name and Description	Design and Statistic of Interest	Cognitive Factors	Operations
1. Coded Messages: Form associations between abstract symbols and corresponding words and combine these symbols into statements.	Introductory Phase: Learn associations of 4 symbols/subset, subset=3. Test Phase: Decide if word message and symbol message are identical. Number decoding symbols (4,5,6) RT and error.	Associative Memory Logical Reasoning or Logical Evaluation, Perceptual Speed.	Basic: Encode words, store their meaning, encode symbols, use visual memory to compare the meanings of the words and symbols, combine these symbols into statements, and execute a proper response.
2. Emergency Procedures: Learn and recall information organized as an ordered set of statements.	Introductory Phase: Learn in order, the first 4, the first 8, and then all 12 steps in a procedure. Test Phase: Decide if stated ordinal relationship between two steps is correct. RT and error.	Does not translate easily into derived Factors.	Encode a series of statements with regard to their serial order, string and retrieve this information, compare serial positions within the stored information, and execute a proper response.
3. Time Check: Read, adjust and compare values from digital and analog information displays.	Introductory Phase: Learn association between clock face and digital display. Learn relationship among the different time zones. Test Phase: Decide if a digital display for a given time zone equals a previously presented clock display for a given time zone. Clock faces rotated (0,90,180, 270^{0}), with outside arrow pointing to clock's 12 o'clock position.	Perceptual Speed, Spatial Relations, Numbericl Computation, Logical Reasoning.	Cognitive: Encode a digital stimulus, store this information, encode an analog or figural stimulus, compare the semantic information from both stimuli, and execute a proper response.
4. Security check: Learn and recall an ordered set of meaningful paired-associates	Introductory Phase: Learn in order, the first 4, the first 8, and then all 12 locations and their associated security designation. Test Phase: Decide if a statement comparing 2 locations in terms of the ordinal relationship of their associated security status is correct. RT and error.	Associative Memory, Logical Reasoning, Sequential Memory.	Encode a series of paired-associates, store and retrieve this information, compare serial positions within the stored information, execute a proper response.
5. Communications Control: Remember hierarchically structured verbal information	Introductory Phase: Learn 3 channels and associated codes, with each code associated with different types of information reflecting traffic flow. Test Phase: Decide if a channel is correctly associated with 2 questions requesting specific information regarding traffic flow. RT and error	Verbal Comprehension, Logical Reasoning Associative Memory.	Cognitive: Encode, store and retrieve linguistically presented information to compare semantic input with this stored information, and to execute a proper response on the basis of the comparison.

Table 3 (Continued)

Air Force Learning Tasks Used to Obtain Learning Curve Parameters

Task Name and Description	Design and Statistic of Interest	Cognitive Factors	Operations
6. Direction Judgment: Store and retrieve information concerning visual forms and their relationships within a two-dimensional array	Introductory Phase: Study the locations of 12 features on a map Test Phase: Decide if a feature is in the suggested direction with regard to a previously presented feature. Orientation of top of previously presented feature (North, South, East, West). RT and error	Visual Memory, Spatial Orientation, Perceptual Speed.	Basic: Encode, store and retrieve visuo-spatial information, to compare perceived information in memory and to execute a proper response based on that comparison process.
7. Strategic Decision: Remember numerical values associated with other stimuli and perform numerical operations on these values.	Introductory Phase: Associate numerical values to a set of stimuli and sum the points of two subtests. Associate one rule to one ratio between the two subsets and another rule to another ratio. Test Phase: Answer (Y/N) if a suggested rule applies to two subsets of stimuli.	Number Facility, General Reasoning, Associative Memory.	Cognitive: Encode, store and retrieve quantitative information to apply computational rules to this information, and to execute a proper response based on the results of the computation.

speed-accuracy tradeoff, signal detection, or as found here, a paired-associate paradigm. These paradigms may allow one to identify critical task components and provide a theoretical framework for identifying behaviors (or processes) of interest.

One potential negative aspect of this experimental approach is the possibly limited transferability of an "experimental" task to a "real world" task. The extent of transferability would need to be empirically determined.

Integration of human performance measurement disciplines. The reorganization of the Air Force's learning parameter tasks into a format depicted by Table 3 suggests that one can integrate different methodologies of human performance measurement. Table 3 shows that one may describe the same behavior in terms of (a) job behaviors that may be used to represent a job-relevant criterion, or (b) abilities that are defined in terms of psychometric-based methodologies (e.g., the factor analytic methodology used to define the cognitive factor of Associative Memory), or (c) the experimental-based methodology that isolates the operation of encoding, comparing, combining, etc.). The column titled "Cognitive Factors" identifies abilities that are described in the psychometric literature, while the column titled "Operations" identifies stages of information processing that are described in the experimental literature.

To the extent that psychometric-based methodologies represent selection measures, experimental methodologies are representative of training approaches, and job-relevant measures represent performance measurement, the mapping represented in Table 3 suggests a possible technique for linking these three human resource management fields.

For example, the table suggests that it is possible that a measure of perceptual speed from the psychometric literature may relate to a criterion task which is based on job analytic methodologies, while a measure of perceptual speed from the experimental literature may relate to both the psychometric measure of perceptual speed and the criterion task. The experimental task may, therefore, be unlikely to exhibit incremental validity in the statistical sense. However, systematic study of the experimental task may highlight the independent variables and their respective levels that have the most marked effect on perceptual speed and criterion task performance. Identification of powerful independent variable effects could then suggest avenues for treatment intervention (i.e., training of relevant behaviors or strategies predictive of performance).

The table also suggests that one may systematically observe a subject's learning of a criterion task. For example, systematic observation of the subject's task learning behaviors may identify differences between successful and less successful subjects in the memory aids they employ, for example, in the Time Check Task (3). Of interest would be the Subject X Learning Parameter interaction for Time Check, which according to the table would be represented by the relationship between perceptual speed (psychometrically defined) and learning parameters, and the relationship

between perceptual speed (experimentally defined) and learning parameters. The intent is to observe if (a) perceptual speed predicts learning performance, (b) components of perceptual speed (experimentally defined) predict learning performance and (c) components of perceptual speed (experimentally defined) predict different components of learning at different stages of task competency. The intent is not to go into an "infinite Do-Loop of research", but rather to suggest that one can systematically map criterion, psychometric literature measures and experimental literature measures to take advantage of the technologies of these respective disciplines.

Measures of Automaticity

Table 4 contains tasks designed to measure individual differences in movement toward automaticity. Automaticity reflects a developed competency in a task due to learning (Schneider, 1985). One question that has been raised regarding automaticity research deals with whether the learning reflects a more specialized way of using the same resources or whether the resources are different from the ones originally used as one learns the task (Logan, 1985). Either strategy could elicit the gradual leveling of performance across stimuli of varying complexity. This effect can be studied in the following three tasks that are employed by the Air Force in its research effort studying automaticity with the intent of isolating those factors accounting for the qualitative changes found with extended practice.

The Perceptual Matching Task (1) represents a scanning task of varying difficulty. A focus of interest would be the effect of practice on individual differences to the extent that the more difficult matrices (e.g., larger and less discriminable) elicit performance (e.g., reaction times) that approach the reaction times of less difficult matrices. (The less difficult matrices would be smaller and more discriminable than their more difficult counterparts.)

The Attribute Comparison Task (2) is a Posner Task, with an embedded variation of type of identity (i.e., physical vs. nominal vs. categorical). A focus of interest would be the effect of practice on individual differences to the extent that more complex matches (e.g., categorical matches) elicit performance that approaches less complex matches (e.g., physical matches).

The Arithmetic Operations Task (3) contains a homogeneous set of operations for all but one of the blocks of trials. A focus of interest could be the effect of practice on individual differences to the extent that more difficult problems (e.g., larger sums) elicit performance that approaches performance for less difficult problems (e.g., smaller sums). Individual differences could be observed for each block of trials.

When observing movement toward automaticity, one may be interested in the interaction between initial task performance and the rate of improvement in task performance. One would then be observing the interaction between initial ability on the task and learning rate for that task.

Table 4

Air Force Tasks Designed to Measures Differences in Movement Toward Automaticity

Task and Description	Design	Statistic of Interest
1. Perceptual Matching: Decide if two matrices match.	Sessions = 8, 96 trials/session Matrix Size (3,5,7,9) elements Different Degrees of mismatch (0,1,2, all elements).	1. Slope and interest of positive and negative decisions. 2. Session x Slope interaction for each decision. Session x Intercept interaction for each decision. 3. Trial effect with slope and intercept. 4. Initial Level x other factor(s) using curve fitting procedures.
2. Attribute Comparison: Decide if two sets of items match	Sessions = 6, 540 trials/sessions. Type of Match Required (Physical, Normal, Categorical) Each type of match as 45 trials/ sessions.	1. Initial Level x other factor(s) using curve fitting procedures.
3. Quantitative Fact Retrieval: Arithmetic Operations	Sessions = 4, 128 or 144 problems/ session. Each session tests a different type of arithmetic operation.	1. For addition, fits RT slope across the different sums. 2. Initial Level x other Factor(s) using curve fitting procedures. 3. Detailed analysis of error data.

Initial ability could also be defined by a classical psychometrically-oriented measure or by a criterion task correlated with the automaticity task of interest. Two relevant questions emerge:

(a) how does performance on the correlated measure interact with initial performance on the automaticity task;

(b) how does psychometric ability or criterion task performance interact with practice on the automaticity task?

These questions could be answered by a split-plot repeated measures analysis, with the Group variable derived by stratification of a psychometric correlate. For example, assuming that a Perceptual Speed measure (developed according to classical psychometric procedures) predicts performance for this task, one can determine how high and low ability subjects on this psychometric measure differ in their respective learning rates for Perceptual Matching (1).

Another strategy could include systematic variation in the introduction of the psychometric and criterion task measures at different stages of practice on the automaticity task. For example, one can systematically introduce a measure of Perceptual Speed or a related criterion task, perhaps one requiring target detection, in one or more of the eight sessions of the Perceptual Matching Task. The goal is to observe how well increased competency on automaticity of the Perceptual Matching transfers to performance in a psychometrically-based ability or criterion measure.

In-House Experiments and Updating of Some Tests

Table 5 is a concatenation of Christal's listing of Air Force experiments performed in-house and the limited listing of the updating of in-house and other tasks (Christal, 1984, 1985). Brief discussions of each grouping of tasks, (e.g., Triad Sub-Battery) follow.

The Triad Sub-Battery employs a paradigm in which all the information for a trial is presented simultaneously. A primary focus is the variation of cognitive and linguistic complexity within the context of a binary classification RT task.

The Speed Level Battery measures speed and level scores in the reasoning, quantitative and verbal domains. This battery contains a number of tasks employed in the Baseline Measures Program (Christal, 1984). The Speed Level tasks that overlap with that program include simple and choice RT (9,10), Sunday-Tuesday Addition (14), Sentence Picture Verification (15), and Number Facts (16). The only difference between the Baseline Measures and the Speed Level Battery appears to be the number of trial blocks and the number of trials per block (see Table 2 for comparison).

Table 5

Air Force In-House Tests

Task Category	Task Name and Description	Task Design	Statistic Of Interest
Triad Sub-Battery	Three stimuli presented simultaneously in Triangular format, with one of the bottom 2 stimuli matching the top stimulus.		
	1. Simple RT - Respond to any light on the screen.		
	2. Choice RT - Press response key under the star, which occurs randomly on the left or right side.		
	3. Physical Identity: CUT CUT SET		
	4. Name Identity: GET set get		
	5. Category Identity: HARP CAN PIANO		
	6. Meaning Identity: RESPITE REVENGE REST		
	7. Sound Identity: SLEIGH SLY SLAY		
	8. Relation Identity: BIRD-FLY SNAKE-SLITHER CAT-MEOW		
Speed Level-(1984) Measure Speed and level scores in the reasoning, quantitative, verbal domains.	9. Simple RT	Blocks = 3	
	10. Choice RT	60 Trials	
	11. Remote Analogies: Verbal anologies	25 Trials	
	12. Number Sets: Select the atypical number series.	20 Trials	
	13. Fact Verification: Determine whether a simple sentence was true or false.	Blocks = 2, 26 Trials/Block Sentence types Noun with characteristic Noun in a category	
	14. Sunday-Tuesday Addition: Perform symbolic arithmetic with Monday = 1 and Sunday = 7. Answers are days of the week not numbers.	35 Trials	
	15. Sentence-Picture Verification: Determine truth or falsity of simple sentences like the following: A B The A is not followed by the B.	Blocks = 4, 16 trials/Blocks Design: "To Follow" vs "To Precede" Active vs. Passive Positive vs. Negative A first vs. B first	
	16. Number Facts: A separate block for each - addition, subtraction, multiplication, division, mixed.	Blocks = 5, 15-16 Trials/Block	

Table 5 (Continued)

Air Force In-House Tests

Task Category	Task Name and Description	Task Design	Statistic Of Interest
	17. Memory Identity: See Meaning Identity in Trial Sub-battery.	Blocks = 3, 28 Trials/Block	
	18. Simple Symbolic Arithmetic: Given A = 6, B = 3, C = 5, C-B = ?	21 Trials	
	19. Complex Symbolic Arithmetic: Given C = B/2, B = 23-17, B A = ?	15 Trials	
	20. CVC - Word Pairs: Memorize 14 associations between trigrams and common nouns and choose synonym of the noun associated with a trigram.	28 Trials, Nouns (moderately easy, moderately difficult)	
	21. Arithmetic Tracking: Perform six successive arithmetic operations.	Blocks = 2, 20 Trials/Block	
	22. Memory Span: Decide if successively presented suggested contiguous numbers were present and contiguous in a prior set.	Blocks = 3, 10 Trials/Block Memory Load Blocked (5, 7, 9 numbers). Number of numbers tested (2 to 5).	
	23. Three-Term Series: Compare and make a choice based on information in two consecutive sentences.	Blocks = 3, 7 items/Block. Variations: 1. markedness (taller, shorter 2. negation in first vs. second clause. 3. congruence in relational adjective in the clause vs. in the question.	
	24. Trait Levels: Associate a trait with a letter, e.g., A = fast, B = slow, C = medium and 2 letters at a time on that trait.	7 items; 6 paired comparison/item	
	25. Rule Application: Based on rules, transform first 2 letters in a string and transform that result with the next letter, etc.	Blocks = 6, 10 Items/Blocks	
Working Memory Battery-Evaluate fixed attentional space shared for short-term storage and execution	26. Revised ABC Test: Associate 3 letters to their respective numeric operations of varying difficulty.	Processing Load (0,1,2 operations). Time in Memory determined by presentation order (1st, 2nd, 3rd) x Test order (1st, 2nd, 3rd).	
	27. Digit Span with Partial Recall: Memorize sequential order of 9 digits.	Equal number of trials for primacy, recently and central effects.	
	28. Missing Digit: Decide which digit is missing.	Memory Load (6, 8, 12)	
	29. Fast Running Recall: Memorize digits of variable length and repeat the last three digits presented.	Presentation Rate (4 or 6 digits per second). Digit Length (4, 8, 12, 16 numbers)	
	30. Letter-Digit: Given a list of digits and associated letters continuously appearing, match the number to its letter.	Blocks = 3, 20 trials/Block List Length (3, 6, 9, digit-letter associations).	Change in median response time over 20 trials
	31. Brown-Peterson a: Memorize 6 digits; respond to varying number of questions regarding a visual display; recall original digit sequence.	Number of intervening questions (4, 8, 16)	

Table 5 (Continued)

Air Force In-House Tests

Task Category	Task Name and Description	Task Design	Statistic Of Interest
	32. Brown-Peterson b: Memorize digits; answer questions regarding the relative order of the digits; recall original sequence.	Memory Load (3 vs. 6 digits) Linguistic difficulty; e.g., 8 does not come before 3 vs. 8 comes before 3.	
	33. Word Recognition Threshold: Recognize common English words.	Word Length (3 vs. 4 letters) Exposure time (17, 33, 50 msec.)	
Spatial-Visualization	34. Figure Synthesis Task: Synthesize line segments and decide if a one or two segment part was contained in the synthesized figure.	Blocks = 4, 24 trials/Block Number of line segments (3, 4, 5, 6).	
	35. Palmer-Figures Part-Whole Verification Task: Decide if lines of a figure were contained in a prior figure.	96 Items Between Ss (Self paced study time of first figure vs. 500 msec. study time vs. 3000 msec. study time) Goodness of Form (closure and proximity values of lines.	
	36-39. Standard Tests-Identical Pictures, Block Synthesis Form Board, Gestalt Completion.		
Short-Term Memory Battery	40a,b. Memory and Visual Search: Decide whether a probe is present or absent.	Memory Search - Memory Load 1, 2, 3, 4, 5 consonants). Visual Search-Display Load (1, 2, 3, 4 consonants).	RT slope regressed on memory; memory load and display load; RT y-intercepts from regression
	41. Peterson and Peterson Task with Subtraction as the intervening task: Memorize a consonant trigram; Subtract a consonant iteratively from a given number; recall trigram.	With Ss (2, 4, 6, 8, 12 interations). Between Ss: Meaningfulness (Word vs. trigram) Difficulty (Subtract 3 vs. subtract 2).	
	42. Iconic Memory: Recall the prompted consonant in a 2x8 matrix of consonants exposed for 50 (msec.).	Time between Matrix disappearance and prompt (50, 100, 150 msec.). Time between prompt and matrix appearance (0, 50, 100 msec.).	
	43. Perceptual Threshold: Identify stimuli presented for brief exposure.	Trials = 150 Exposure Time (17, 33, 50 msec.) Type of stimuli (upper and lower case letters digits). Stimulus length (Single vs. double digits, 3 vs. 4 letter words).	
Verbal Learning	44. Paired-Associates Learning with Imagery Systematically Varied: Learn 10 word pairs for each trial. Answer 10 multiple choice questions testing memory for word pairs, with only the first and last letters of the associated word and distractors appearing.	Trials = 12; 10 questions/trial. Within Ss: Imagery of pairs (low, medium, high) Trials (4 trials/imagery level) Between Ss: Type of training (No training vs. semantic elaboration on last 6 trials vs. repetition on last trials).	
	45. Analogical Reasoning with Simple Recall: Present analogy to be completed based on prior analogy. After 20 trials, prompt S to remember one of the prior analogies.	RT for A: B and for C:_.	

Table 5 (Continued)

Air Force In-House Tests

Task Category	Task Name and Description	Task Design	Statistic Of Interest
	46. Self-pased Paired Associates Learning: Think of a sentence to connect a word-pair; recall the sentence when presented with half of the word pair.		
Miscellaneous (1984)	47. Numerical Operations: Computerized version of test used by the Air Force in the 1950's.		
	48. Memory for Lists in Different Spatial Arrangements: Memorize 12-item list.	Trials = 36. Array Arrangement (Clock-like, ladder, step-like). Structure of Prompt (Pictorial, i.e., lists's structure maintained except for blank to be recalled stimulus vs. verbal, i.e., "Recall what item fell in a given position number").	
	49. Alphanumeric Series Matching: Decide if two side-by-side letter strings are same or different.	Trials = 72. Display Load (8, 9 characters)	
	50. Picture Paired Associates Learning: Memorize 10 picture pairs and select the correct associate in a multiple choice format.	Sets = 3; 10 pairs/set	
	51. Nonsense-Picture Paired Associates: Same as above, but nonsense picture used.		
	52. Trigram Paired Associates Learning: Same as picture pair, but manipulate meaningfulness.	Sets = 6, 10 pairs/set Meaningfulness (3 high vs. 3 low meaningful sets of trigrams)	
	53. Word Paired Associates Learning: See #50.		
	54. Reading Rate: Read and memorize comprehension of six sentences presented one at a time.	Passages = 6, 8-10 questions/passage.	
	55. Free Recall: Memorize 70 words for each of 10 categories.	Categories = 10 Non-referenced word frequency (high, medium, low).	
	56. Memory for Classes: Memorize 45 sets of 5 objects for later recall.	Trials = 45; Set Manipulation (Test object presented for memory vs. test object not presented but set member of object vs. test object not presented).	
	57. Memory for Digit Order: Decide if suggested 7-string digit is in the same order as a prior string.	Trials = 80. Order manipulation (same order vs. transposition of adjacent digits).	
	58. Perceptual Threshold: Identify stimuli presented for brief exposure.	Exposure Time (16.2, 33, 50 msec.) Type of stimuli (upper and lower case letters, digits) stimulus length (single vs. double digits, 3 vs. 4 letter words).	

The Working Memory Battery contains tasks requiring continual updating of information (27,29,30), arithmetic operations (26), memory of a large number of stimuli (27,28), memory with intervening interference tasks (31, 32) and perceptual recognition (33). These tasks appear to require increased effort as defined by the magnitude of the number of stimuli on which the subject operates.

Spatial Visualization tasks appear to require updating and integration of past and present information (34,35). This battery also contains a number of standardized tests (36-39).

The Short-Term Memory Battery contains the Sternberg task with variations of memory load (40a) and display load (40b). The tasks also represent various stages of visual memory (43,42,40a,40b) and a paired associates task with an extra retehtion component (41). These tasks do not appear to concentrate on difficult stimulus parameters as found for the Working Memory Battery. One may posit that the Short-Term Memory Battery is designed to capture various search behaviors while the Working Memory Battery is designed to capture recursive behaviors and/or behaviors requiring increased effort.

The Verbal Learning Battery employs paired-associate tasks with embedded independent variables (44) and an analogical reasoning task popularized by Robert Sternberg (45).

The Miscellaneous Tasks include various speed level tasks (47,48,49), paired-associates tasks (50-53,55,56), comprehension (54), serial order memory (57) and perceptual threshold (58).

The tasks listed in Table 5, like those in Table 2, contain more trials per task condition than do Project A tasks. Table 5 does not contain reliability information as reported in Table 2. There is overlap in the content of the tasks listed in Tables 5 (Air Force in-house tests) and Table 2 (Air Force Baseline Tasks).

NAVAL SPATIAL ABILITY TESTS

Table 6 tasks are being developed by the Navy (Christal, 1984, 1985). These tasks are included in the present review because of their emphasis on spatial abilities, a primary focus of Project A measures. The Navy's tasks are categorized as those involving static displays, displays with moving elements and spatial orientation.

The static displays include the Shepard Mental Rotation task (2) and SAME-DIFFERENT judgment task of complex and perturbed stimuli (Cooper, 1976; 1980). Also included are tasks (3,4,5,6) that capture different aspects of Kosslyn's theory of mental imagery (Kosslyn, 1981).

The moving display tasks appear to stress the ability to identify and anticipate intersection points of a trajectory (7,8,9).

Table 6

Naval Spatial Ability Tasks

Task Category	Task Name and Description	Task Design	Statistic of Interest
Tasks Involving Static Displays	1. Figural Comparison: Decide if two simultaneously presented polygons are the same or different (i.e., Perceptual Speed).	Trials = 120; 240 Complexity (No. geometric points) Perturbation (Extent of difference)	Average RT, Errors
	2. Mental Rotation: Decide if two simultaneously presented polygons are the same or different.	Trials = 200; Angular disparity ($0-180^{\circ}$) Match Type (Same or Different).	RT slope regressed across angularity. Intercept of RT. Separate Slope and Intercept for each Decision.
	3. Adding Detail to an Image: Observe a shape repeatedly, with each presentation containing the shape and 1 additional detail. Decide in multiple choice format which shape contains all the details.	Trials = 32-40; Number of successive presentations (3 to 6).	
	4. Form Board	Trials = 240, Vary: 1. No. of Stimulus Elements. 2. Manipulation of Elements for Assembling.	RT estimates for motor time, encoding and comparison time, search time, Rotation Rate, Accuracy.
	5. Image Integration: Mentally integrate array of irregular shapes and then decide in multiple choice format the correct shape.	Trials = 20 to 30; Number of Shapes (2 to 6).	Time to perform integration. Time to decide correct shape. Accuracy.
	6. Surface Development: Decide if, e.g., 2 edges of a flat unfolded cube would fold into an identical cube as a previously presented unfolded cube.	Trials = 72 to 90; No. of manipulated surfaces (1 to 9).	Time to mentally fold first first cube. Time to decide match. Accuracy.
Tests Involving Moving Elements	7. One Moving Element: Decide the intercept point of a trajectory or when the trajectory will reach a point or both.		
	8. Two moving Elements: Trajectory paradigm as #7.	Vary: 1. Type of trajectory (straight vs. curved) 2. Angle of intersection 3. Speed and acceleration 4. Extent of motion extrapolation	
	9. Pseudo Driver Displays: 3-D perspective for trajectory test.	Vary: 1. See Two Moving Elements (#8) 2. Subject's 3D perspective of moving objects	
	10. Geographic Orientation: Simulated route learning task in which Ss are shown a changing display representing movement through an environment.	Vary: 1. Environment complexity 2. Changes in Direction 3. Distance traveled	
Spatial Orientation	11. Spatial Orientation: From within a windowless room in a familiar setting (i.e., college campus) perform different tasks.	Tasks: 1. Point to college landmark 2. Mark direction on paper 3. Draw a campus map	
	12. Memory for Position: Decide if a sequence of lights on a 3x2 grid was presented previously.	Memory Load (2, 3, or 4 lights in sequence)	
	13. Indication of location of moving object: Decide bearing of 3D object after it moves off the screen. Decide time 3D object would be alongside self.		Accuracy of pointing
	14. Geographic Orientation: Draw map or simulated environment.	See Task #10.	Accuracy of map

The table also includes tasks that appear to stress a subject's ability to place himself/herself in a static or moving environment (10,11,13,14). Tasks employing this ability but adapting it to 'real world' environmental cues are included (10,14) as are tasks dealing with updating analog information (12,13). A criterion task of orientation ability contains subtasks of orienting oneself in a familiar environment by pointing to directions of well-known features and drawing a map that contains those features (11).

The tasks listed in Table 6 were selected and developed under contract with Earl Hunt and J. W. Pellegrino.

ARMY'S PROJECT A PREDICTOR BATTERY

The Project A predictor battery is designed to be an expansion of the number of constructs that may be used as predictors to select and classify Army recruits into different military occupational specialties (Peterson, 1985). A brief description of Project A measures and the role that the developing cognitive and psychomotor measures play in the project was presented in the present report's Introduction section. The report will now describe those measures.

Table 7 contains the Project A computer and paper-and-pencil tests. The tests represent some of the hypothesized dimensions of primary importance for performance in the Army and are designed to discriminate reliably among the ability levels of those recruits now entering the Army. The reader is referred to Eaton et al. (1985) or Peterson (1985) for further discussion of Project A predictors on topics not included in the present report.

Some of the tests listed are not currently considered as necessary as they once were. Necessity may be determined by construct uniqueness and fitting test administration constraints. For example, the Shapes Test which is designed to measure field dependence is not currently being considered as part of the battery. Support for this decision may be found in the factor analytic studies showing that field dependence and tests similar to Assemble Objects, another test in the Project A battery, consistently load on the same factor (Mos, Wardell, & Royce, 1974). The purpose of including the Shapes Test and other tests <u>not</u> currently part of the battery is to maximize possible comparisons between Project A developed tests and the tests developed by the other services. Other tests no longer in the battery include the Path, Reasoning 2 and Orientation 1 Tests. The Orientation 3 Test is currently called the Map Test.

Comparisons Between Project A Tests and Air Force and Navy Tests

A primary purpose of this report is captured by this section. The section contains tables that cross-reference Project A tests with the Air Force and Navy tests reviewed earlier in the report. Tables 8, 9, 10, 11, and 12 represent a 'cross-matrix' of the Project A tests with the AFHRL programs of 1984 and 1985 and the Navy's 1985 program.

Table 7

The Army's Project A Cognitive and Computer Predictor Battery

Construct and Task Name	Task Description	Design	Statistic of Interest
Spatial Visualization Rotation			
1. Assembling Objects	Decide how an object looks when parts are put back together.	Items = 40; Label (Label contiguous parts vs. don't label contiguous parts)	% Correct
2. Object Rotation	Decide whether two figures are the SAME or DIFFERENT.	Items = 90; Orientation (Rotation vs. Flipped Over)	% Correct
Field Independence			
3. Shapes	Decide which simple figure is embedded in a more complex figure.	Items = 54	% Correct
Scanning			
4. Path	Decide which of a given number of defined paths is the shortest between 2 points.	Items = 44	% Correct
5. Mazes	Decide which maze entrance passes through the maze and to one of its exit points.	Items = 24	% Correct
Induction			
6. Reasoning 1	Identify pattern among figures and decide next figure in series.	Items = 30	% Correct
7. Reasoning 2	Identify the anomalous figure.	Items = 32	% Correct
Spatial Orientation			
8. Orientation 1	Given a circle with a labeled North heading, decide which of 5 other heading-labelled circles are reoriented consistently.	Items = 150; 5 items/set	% Correct
9. Orientation 2	Mentally rotate a frame around to the bottom of a scene, and decide what the orientation of a feature in the frame would be.	Items = 20	% Correct
10. Orientation 3	Given the direction of a landmark, decide the direction of travel to other landmarks.	Items = 24	% Correct
Cognitive/Perceptual Tests:			
11. Simple Reaction Time		Trial = 10	$\bar{X}$, S.D., % Errors
12. Choice Reaction Time		Trial = 30	$\bar{X}$, S.D., % Errors

Table 7 (Continued)

The Army's Project A Cognitive and Computer Predictor Battery

Construct and Task Name	Task Description	Design	Statistic of Interest
13. Perceptual Speed and Accuracy	Decide if two stimulus arrays match.	Trial = 36; (alpha, numeric, symbolic) x 3 (2,5,9 Characters) x 2 (SAME vs. DIFFERENT Judgments).	$\bar{X}$, S.D., Slope Intercept, % Errors
14. Target Identity	Match a military vehicle or aircraft target to one of 3 possible targets.	Trial = 36, Agularity x Discriminability	$\bar{X}$, S.D., Slope Intercept, % Errors
15. Short-Term Memory	Sternberg Memory Scanning Task	Trial = 36; 3 (1,3,5 memory load) x 2 (letters vs. symbols) x 2 (Presence vs. Absence Judgments)	$\bar{X}$, S.D., Slope Intercept, % Errors
16. Number Memory	Continually perform numerical operations on a sequentially presented set of numbers.	Trial = 26 or 18	$\bar{X}$, S.D., Slope Intercept, % Errors
17. Cannon Shoot	Rendezvous with a target by firing a shell from a stationary position.	Trial = 36; 4 (Right, Down, Left, Up Direction to fire)	$\bar{X}$ Time, X Log (distance +1) Error
Psychomotor Tests			
18. Target Track 1	Rendezvous with and remain on top of a target moving along a marked path by moving a joystick.	Trial = 18, 3 (max. crosshair speed) x 3 (Difference btwn Speed of Target and Max. Crosshair Speed) x 3 (No. of turns in path)	$\bar{X}$ log (distance +1 errors
19. Target Track 2	Rendezvous with and remain on top of a target moving along a marked path by moving a horizontal and a vertical slide.	Trial = 18, 3 (Max. crosshair speed) x 3 (Difference btwn Speed of Target and Max. Crosshair Speed) x 3 (No. of turns in path)	
20. Target Shoot	Rendezvous with a target that is unpredictably changing directions by moving a joystick and pressing a button to fire.	Trial = 30, Speed of Target x Max Speed of Crosshair x Mean Length of each Segment traversed by the target.	$\bar{X}$ Time to fire Percent Hits

Note. Reaction time data are broken down into Decision, Movement and Total Time.

Reaction time data are presented in both raw form and by log transformation, with the fastest and slowest RT trimmed off for each stage.

The independent variable information is taken from McHenry & McCue (1985) and Toquam, et. al. (1985).

Table 8

Air Force and Navy Tests Similar to Project A S
patial Visualization Tests

		Project A Spatial Visualization Tests				
		Rotation		Field Independence	Scanning	
Table Number	Test Program/Battery	Assemble Objects	Object Rotation	Shapes	Path	Maze
Air Force:						
1	BAT		Mental Rotation(5)	Embedded Figures(10)		
2	Baseline	Ichakawa short term Memory for Position(27)	Rotated Figures(25)			
3	Learn Curve					
4	Automaticity					
5	Speed Level					
5	Miscellaneous					
5	Verbal Learning					
5	TRIAD					
5	S.T.M.					
5	Working Memory					
5	Spat. Vis	Figure Synthesis(34) Block Design(38)		Palmer-Figures(35)		
Navy:						
6	Static Displays	Image Integration(5) Form Board(4)	Mental Rotation(2)	Form Board(4)		
6	Moving Displays					
6	Spatial Orient.					

Note. B.A.T. = Basic Attributes Test Program
Learn Curve = Tests designed to capture learing parameters
TRIAD = Tests in which standard stimulus and the two choices form a triangle
S.T.M = Short-Term Memory
Spat. Vis. = Spatial Visualization
Spat. Orient. = Spatial Orientation

Table 9

Air Force and Navy Tests Similar Project A Reasoning and Spatial Orientation Tests

		Inductive Reasoning		Spatial Orientation		
Table Number	Table Grouping	1	2	1	2	3
Air Force:						
1	B.A.T.			Mental Rotation (5)		
2	Baseline	Last Letter of Pattern (20); Three Term Series (16)	Patterns of Letter Series (18); Three Term Series (16)	Rotated Figures (25)	Rotated Figures (25)	
3	Learn Curve			Time Check (partly) (3)		Direction Judgment(6)
4	Automaticity					
5	Speed Level	Three Term Series (23)	Number Sets (16); Three Term Series (23)			
5	Miscellaneous					Memory for lists in different spatial arrangements(48)
5	Verbal Learning	Analogical Reasoning (45)				
5	TRIAD	Relational Identity (8)				
5	S.T.M.					
5	Working Memory					
5	Spat. Vis.					
Navy:						
6	Static Displays			Mental Rotation (2)	Mental Rotation (2)	
6	Moving Displays					
6	Spatial Orient.			Pseudo Driver (9) Indicate Location of Moving Target (13)	Pseudo Driver (9) Indicate Location of Moving Target (13)	Spatial Orient, (11) Geographic Orient. (10)

Note. B.A.T. = Basic Attributes Test Program; Learn Curve = Tests designed to capture learning parameters; TRIAD = tests in which standard stimulus and the two choices form a triangle; S.T.M. = Short-Term Memory; Spat. Vis. = Spatial Visualization; Spat. Orient. = Spatial Orientation.

Table 10

Air Force and Navy Tests similar to Project A Reaction Time and Perceptual Speed and Accuracy Tests

Table Number	Grouping	Reaction Time: Simple	Reaction Time: Choice	Perceptual Speed & Accuracy (PS&A): PS&A	Perceptual Speed & Accuracy (PS&A): Target Identity
Air Force:					
1	B.A.T.		Decision-Making Speed (8) (Partly)	Encoding Speed (4)	Mental Rotation (5)
2	Baseline	Simple RT(1)	Choice RT(2)	Posner Physical Identity (3)	Rotated Figures (25)
3	Learn Curve				
4	Automaticity			Attribute Comparison (2) Perceptual Matching (1)	
5	Speed Level	Simple RT (9)	Choice RT (10)		
5	*Miscellaneous*			Alphanumeric Series Match (49)	
5	Verbal Learning				
	TRIAD	Simple RT (1)	Choice Rt (2)	Posner Physical & Name Ident. (3, 4)	
5	S.T.M.				
5	Working Memory				
5	Spat. Vis.				
Navy:					
6	Static Displays			Figural Comparison (1)	Figural Comparison (1) Mental Rotation (2)
6	Moving Displays				
6	Spatial Orient.				

Note. B.A.T. = Basic Attributes Test Program; Learn Curve = Tests designed to capture learning parameters; TRIAD = Tests in which standard stimulus and the two choices form a triangle; S.T.M. = Short-Term Memory; Spat. Vis. = Spatial Visualization; Spat. Orient. = Spatial Orientation.

Table 11

Air Force and Navy Tests Similar to Project A Short-Term Memory, Number Memory and Movement Judgment Tests

Table Number	Table Grouping	Memory: S.T.M.	Memory: Number Memory	Movement Judgment: Cannon Shoot
Air Force:				
1	B.A.T.	Item Recognition (6)		
2	Baseline	Sternberg No. (7a), Words (7b); Visual Scan (21)	Numerical Operations (8)	
3	Learn Curve		Strategic Decision (7)	
4	Automaticity		Quantitative Fact Retrieval (3)	
5	Speed Level		Arithmetic Tracking (21); Simple Symbolic Arithmetic (18); Number Facts (16)	
5	Miscellaneous		Computerized Numerical Operations (47)	
5	Verbal Learning			
5	TRIAD			
5	S.T.M.	Memory & Visual Search (40a, b)		
5	Working Memory		Revised ABC Test (26)	
5	Spat. Vis.			
Navy:				
6	Static Displays			
6	Moving Displays			One Moving Element (7); Two Moving Elements (8); Pseudo Driver Task (9)
6	Spatial Orient.	Memory for Position (2)		Indicate Location of Moving Target (13)

Note. B.A.T. = Basic Attributes Test Program; Learn Curve = Tests designed to capture learning parameters; TRIAD= Tests in which standard stimulus and the two choices form a triangle; S.T.M. = Short-Term Memory; Spat. Vis. = spatial Visualization; Spat. Orient. = Spatial Orientation.

Table 12

Air Force and Navy Tests Similar to Project A Psychomotor Tests

Table Number	Table Grouping	Psychomotor: Target Track 1	Target Track 2	Target Shoot
Air Force:				
1	B.A.T.	Time Sharing (3)	Time Sharing (3) 2-Hand Coordination (16)	Time Sharing (3)
2	Baseline			
3	Learn Curve			
4	Automaticity			
5	Speed Level			
5	Miscellaneous			
5	Verbal Learning			
5	TRIAD			
5	S.T.M.			
5	Working Memory			
5	Spat. Vis.			
Navy:				
6	Static Displays			
6	Moving Displays	Two Moving Objects (8)		Two Moving Objects (8)
6	Spatial Orient.			

Note. B.A.T. = Basic Attributes Test Program; Learn Curve = Tests designed to capture learning parameters; TRIAD = Tests in which standard stimulus and the two choices form a triangle; S.T.M. = Short-Term Memory; Spat. Vis. = Spatial Visualization; Spat. Orient. = Spatial Orientation.

The tables list the names of the programs and their respective tests which appear to be similar to Project A tests. For example, the Form Board Test of the Navy's Static Display Test Battery appears analogous to Project A's Assemble Objects Test.

Although brief explanations for the suggested overlaps between Project A tests and the Air Force and Navy tests follow, this author is aware of the limited reliability of this "sorting" task with an n=1. The reader may refer to Table 7 for descriptions of the Project A tests and Tables 1 to 6 for descriptions of the other tests to decide whether he/she agrees with the proposed overlaps/relationships listed in Tables 8-12.

The reader is encouraged to perform two "search" tasks as he/she is reading the upcoming text. First, he/she may look at the relevant "crossmatrix" table to get a sense of the other tests to which a Project A test is suggested to be analogous. Second, the reader may use the crossmatrix table to refer to the table number in which the task suggested to be analogous to a Project A test is described.

For example, the text suggests that the Image Integration Task appears analogous to the Project A Assemble Objects Tests. Table 8 contains the column heading "Assemble Objects" under which suggested analogous tasks are listed. Table 8 shows that one suggested analogous test to Assemble Objects, the Image Integration Task, is part of the Navy's Static Display Battery and directs the reader to Table 6, test number 5, to obtain more information regarding the Image Integration Task. The text just prior to Table 6 contains a brief description of the test and the battery, as well. These and other comparisons between each Project A test and the Air Force and Navy tasks follow.

<u>Assemble Objects (Table 8)</u>. The Form Board and Image Integration Tests require that a subject put together parts of a geometric figure to form that figure. This appears to be directly analogous to the Project A Assemble Objects Test, except that Form Board and Image Integration are computer administered while Assemble Objects is in booklet form.

The Figure Synthesis and Ichakawa tests require that the subject integrate the successive presentation of different parts of a figure and to assemble the parts together to form a whole figure.

<u>Object Rotation (Table 8) and Orientation (Table 9) Tasks</u>. Spatial tests may require different types of distinctions. These include distinguishing between:

(a) a matching figure <u>versus</u> its mirror image (rotated or unrotated);
(b) a rotated matching figure <u>versus</u> its mirror image or a different figure;
(c) a matching figure <u>versus</u> a different figure;
(d) a matching trajectory <u>versus</u> a different trajectory.

The Project A Object Rotation Task requires discrimination between a figure and its mirror image. The BAT's Mental Rotation Task also contains trials in which a figure's mirror image is rotated. However, this Mental Rotation Task requires that the subject decide that a mirror image is a match, while Project A's Object Rotation Test requires that the subject decide that a mirror image is a mismatch. The other tests listed under Object Rotation do not require distinguishing between an object and its mirror image.

The Spatial Orientation task requires that one identify the correct orientation of an object, once having mentally rotated the object to a prespecified position.

One may consider identification of a trajectory as a component of a spatial orientation task. For example, when a subject is deciding how an object appears when it is rotated one may argue that the subject is attempting to identify an object when its trajectory comprises its swiveling, or re-orienting, on itself. In other words, while identifying the "swivel" trajectory, the subject matches up parts of the original figure with the swiveled figure. Such reasoning suggests that the Naval Trajectory Tasks (Indicate Location of Moving Target, Pseudo Driver) could relate to Spatial Orientation tasks and that a "swivel" trajectory identification task could relate to the Object Rotation Test.

Also of interest are two analog Air Force tasks, Time Check and Direction Judgment. The Direction Judgment Task requires that one remember the spatial positions of different features on a map, which is similar to Project A's Orientation 3 Task. A component of the Time Check Task requires that one determine the time of a clock face that has been rotated, a task similar to Project A's Orientation 1 Task.

One possible implication of these Air Force tasks is that one can isolate critical parts of job behaviors and use those components as predictors of task performance. One can also embed experimentally manipulated variables within components of a criterion task that is composed of job-relevant behaviors. For example, The Memory for Lists in Different Spatial Arrangements Task (see Table 3) requires that one remember the relative position of various features within a geometric figure. The task is also designed to assess the effect of visually-cued information. The visual presentation mode is represented by the reintroduction of the same figure with the critical feature missing, with the requirement that the subject recall the missing feature. Alternatively, the subject is asked in sentence format to recall the feature that was in a specific position. Comparison between the visually-cued format versus the sentence-format allows for study of the regeneration of stimuli via the visual versus verbal stimulus modes.

Field Independence (Table 8). The Project A Shapes Test is designed as an analog to the Hidden Figures Test, a measure of field dependence. The Air Force BAT program presents two complex figures and requires the subject to decide which of the two figures contains a specific simple figure. Other

variations of field dependent measures may be deciding whether line segments are embedded within a previously presented figure (Palmer Figures) or carefully matching two similar complex figures to see if they totally match (Figural Comparison). A Form Board Test, which operationally requires assembling simple objects into a more complex object, is included here because Form Board Measures typically load on the same Factor as field independence measures. This has been labeled the Flexibility of Closure Factor (Mos, et al.).

Scanning Tasks (Tables 8). This report did not identify analogs of the Project A scanning tasks in the Air Force and Naval Programs that it reviewed.

Inductive Reasoning (Table 9). Inductive Reasoning may require identification of an anomalous series (Project A's Reasoning Test 2 or other tests, such as Patterns in Letter Series or Number Sets). It may also require identification of a given pattern (Last Letter of Pattern Test or Project A's Reasoning Test 1). One may also expect relationships between Reasoning Test 1 with Three Term Series, Analogical Reasoning and Relational Identity Tests.

Simple and Choice Reaction Time (Table 10). The Project A simple reaction time (SRT) task requires that the subject respond to the presence of a stimulus; the task appears analogous to other SRT tasks. The Project A choice reaction time (CRT) task requires that the subject make one response to one stimulus and another response to another stimulus. The Project A CRT task contains an interval between the initial stimulus and the stimulus to which the subject responds. The Air Force Triad CRT task simultaneously presents the "initial" stimulus with the two choices underneath, with all three stimuli forming a triangular pattern. The subject decides which of the two bottom choices matches the top stimulus. The easy condition in the BAT Decision-Making Task, as well as, the Baseline and Speed Level CRT tasks appear analogous to the Project A CRT task.

Perceptual Speed and Accuracy (Table 10). One may describe this task as a SAME-DIFFERENT judgment task or a computer-administered clerical test. By incorporating this Project A task within the SAME-DIFFERENT paradigm, one may conceptualize this task as a variant of either the Posner task (which requires distinguishing between nominal, physical and categorical identity), or the Figural Comparison task (which requires identification of geometric figures with various degrees of similarity) or the Encoding Task (which is a Posner task by a different name). One may hypothesize that the most analogous trials of the Figural Comparison Task to the Perceptual Speed and Accuracy Task would be the Figural Comparison trials that require detailed search rather than search for readily discriminable differences. One may hypothesize similar relationships between Attribute Comparison and Perceptual Matching Tasks with the Perceptual Speed and Accuracy Task. The most analogous task to Perceptual Speed and Accuracy appears to be the Alphanumeric Series Matching Task.

One possible source of confusion is the BAT's Perceptual Speed and Accuracy Task. This task actually requires a motoric response to each symbol in a series, rather than the scanning of two series of symbols to decide whether or not they match.

Target Identity (Table 10). Tests suggested as analogs to the Target Identity Task are those that require comparison of geometric figures of varying difficulty (*Figural Comparison*) and identification tasks with an angular rotation independent variable embedded within the task (Mental Rotation, Rotated Figures).

Short-Term Memory (Table 11). The short-term memory task is used in various programs. The task may use numbers or letters as stimuli. Project A varies memory load as does the BAT program, while the other Air Force programs vary memory and display loads. Neisser's task requiring search for "presence" is also listed here. The Memory for Position Test employs memorization of a sequence of lights as the stimulus memory load.

Number Memory Test (Table 11). Analogs of Project A's Number Memory Test come in various forms. The most analogous appears to be the Arithmetic Tracking Task. The other tests listed require associating a letter with a value and performing numerical operations with the letters. A slight variation of the number-letter association is the association of geometric figures (e.g., bombers) with certain values and determining the ratio of the values for oneself and one's "opponent". This latter task is called the Strategic Decision Test.

Cannon Shoot Task (Table 11). The suggested Cannon Shoot analogs appear to require that the subject determine a trajectory and/or its point of intersection. The tasks may have the target move or the target and the figure firing the trajectory move.

Psychomotor Task (Table 12). Analogs of the Project A psychomotor tasks require maintaining a marker position while performing a compensatory task of varying difficulty (Time Sharing). Also included are two-handed coordination tasks.

A Summary of Similarities Between Project A and the Other Services

A number of predictors used by Project A are, thus, analogous to tests being developed by the other services. These "common" predictors include Assemble Objects, Number Memory, Memory Scanning, Rotation and Target Identity, Reasoning Tests, Perceptual Speed and Accuracy, and Cannon Shoot and Psychomotor Tests. The following brief descriptions identify those constructs that were suggested in the previous section to be found among both to the Project A tests and the Air Force or Navy tests described in this report.

The Number Memory test requires updating information through arithmetic computation. Memory Scanning captures subject scanning strategy for information in short-term memory. The Assemble Objects Test looks at

components of visual memory which require manipulation of information, such as assembling and reorienting parts of a figure to form a whole figure.

Other tests which look at visual memory characteristics include the Object Rotation, Target Identity, and Perceptual Speed and Accuracy Tests. These three tasks may also be described as SAME-DIFFERENT Judgment Tasks, a subject area within experimental psychology containing formal models that allow for predictions of performance patterns across the stimulus parameters within these tasks.

Reasoning tests, although being employed in various forms, essentially require detection of a pattern among stimuli and/or determining the next stimulus for that pattern. This type of behavior has been extensively studied in the concept formation literature, with its formal models predicting performance across stimulus parameters.

Finally, tests similar to the Project A Cannon Shoot and Psychomotor Tasks require anticipating and/or intersecting with a point in a trajectory. These type of tasks appear to differ by (a) the type of moving objects (target versus object that is firing), (b) amount of moving objects, and (c) the coordination of responses required to execute the anticipation and intersecting behaviors (joy stick versus a mechanical slide; continual following of target versus firing from a fixed position).

Differences Between Project A and Other Services

The Project A measures do not include tasks employing the paired-associates paradigm, which allows for the isolation of the acquisition, retention and retrieval stages of learning. The Air Force appears to be making extensive use of this paradigm by manipulating variables that differ in memory load and, visual and linguistic complexity, and observing the effect of these variables on individual differences in learning various tasks. These tasks appear to include components of job-relevant behaviors that have been modified to allow for systematic study of individual differences by this paradigm. One cannot state that, based on the descriptions in Christal (1984, 1985), the component behaviors were identified by formal job sampling techniques. One may, instead, state that the behaviors may be representing hypothesized components of a job.

Three test-administration areas in which the Project A computer tasks appear to be different than the other service's computer tasks are in number of trials per task, dependent measures of interest and resolution of reaction time measurement. Project A computer tasks have fewer trials per task than are found for analogous tasks employed by the other services. Examples of some of these differences are noted in the discussion on the Air Force Baseline Tests. One may pose the question of whether the Project A computer tests are currently capturing acquisition or stabilized performance measures. One may also argue that the other services are capturing data that provide a more detailed look at the parameter differences within each of these experimental tasks. At issue is whether these differences are job relevant.

Interservice differences are found in the measurement of reaction time as a dependent measure. Project A separates a subject's reaction time into Decision and Movement Times. Decision Time represents the interval between onset of the stimulus to which the subject responds and the time the subject initiates the response. Movement Time is the time from response initiation to completion of the response. The other research programs do not appear to make the Decision and Movement Time distinction. Rather, they measure Reaction Time as the subject's total time from stimulus onset to pressing the button. Jensen is a proponent of the Decision and Movement Time distinction (Jensen & Munro, 1979) while Longstreth argues against it (Longstreth, 1984).

A related issue for the measurement of reaction time is the assumption of a strong association between the stimulus and its associated response. That is, one would not want to confound measurement of a subject's decision to respond to a stimulus with the subject's "figuring out" the correct response button to press to represent his/her decision. Standardization of response panel design and of the necessary number of warm-up trials to insure sufficient stimulus-response association appear as readily approachable areas for interservice cooperation. However, the theoretical arguments regarding the "purity" of the Decision Time measurement would still remain.

The third test-administration issue concerns the sensitivity of measurement of a subject's performance (McHenry & McGue, 1985; Rosse, 1985). The Air Force employs a Perceptual Threshold Task that presents stimuli for only 17 msec., suggesting that the Air Force employs sophisticated computer programs with a 1 msec. resolution. The Project A Computer battery has a 10 msec. resolution (C. B. Walker, personal communication, 1986).

As one reviews this report's extensive list of computer tests, one realizes that the use of a microcomputer as an instrument to test the competency of a cognitive skill represents a technological breakthrough in behavioral assessment and skills training. However, technical problems exist. The more dramatic technical problems for consideration when using the microcomputer as a scientific instrument include the use of software that allows for sensitive stimulus presentation parameters and sensitive measurement of subject behavior. The sensitive unit of measurement in reaction time tasks is typically 1 msec. Computer programs having millisecond accuracy have been written that allow for such on/off stimulus presentation and subject measurement (Dlhopolsky, 1982). Identifying how millisecond resolution, response panel design and other problems (Lincoln & Lane, 1980; Rosse, 1985) are being solved provide an avenue for cooperation in test development among the services.

One other difference between Project A and the other services is Project A's dropping the Shapes test from the predictor battery. For example, Table 8 describes an embedded figures task that is used by the Air Force that is analogous to the Shapes test. The Shapes test was originally part of the Project A battery but was dropped because of time constraints in test administration and because it was felt that other tests such as the Assemble Objects Test provide sufficient overlap with the ability tapped by the

Shapes Test. One may still want to keep the Shapes Test on the "back burner." Forms of Embedded or Hidden Figures tests have strong relationships with target detection and identification tasks (Bone, 1978; Thornton, Barret & Davis, 1968), automobile driving behaviors (Mihal & Barrett, 1976), and other tasks or situations that would be of interest to the Army and the other services (cf., Long, 1972).

Again, the other services' research programs covered in this report do not appear to be employing analogs to the Project A Scanning Tasks.

Potential Predictors Not Included in Any Service Battery

Measures of selective attention or dichotic listening (e.g., Gopher & Kahneman, 1971; Gopher, 1982) do not appear that popular, although other paradigms reflecting channel capacity or attentional demand models are employed. Examples include the BAT's Time Sharing Task and the Air Force's Working Memory Battery. One example of a dichotic listening task is found in the Tri-Services Test Battery. This battery was not reviewed by this report, although Christal (1985) does list the battery, with brief descriptions of each test.

The low emphasis that the services appear to be giving channel capacity measures, as in the Broadbent Model (c.f., Broadbent, 1958; Kahneman, 1973) when selecting tasks to represent components of an information processing model should be recognized. Such a shift in the definition of information processing may be resulting in test batteries that are not taking full advantage of certain potential measures, which to paraphrase the intent of the Air Force's BAT program, require that the subject quickly sort, prioritize and act on a continual stream of visual, auditory and tactile information.

One may argue that the Project A Number Memory and Psychomotor Tests represent tasks requiring updating and responding to a continual stream of information. However, the abilities of quickly sorting and prioritizing information do not appear to be as obvious for those tasks. Perhaps, Project A measures may not have been designed to capture these "sorting" abilities because these abilities may not likely show predictive validity for Army jobs.

The Annett Handedness Scale (Annett, 1970), another unique measure not discussed in the present report is a 12-item self-report questionnaire being employed by the United Kingdom (Christal, 1985). The scale allows for a continuous (as opposed to dichotomous) measure of handedness, an indirect measure of laterality (Hardyk & Petrinovich, 1977). The laterality construct has been shown to be related to spatial abilities, which appear to be given a strong emphasis by the Project A predictor battery. One possible relationship of interest is that between laterality and a criterion task predicted by spatial abilities. Such a relationship may carry implications for equipment design and training.

CONCLUDING COMMENTS

It appears that the services are identifying profiles of ability competencies and critical job behaviors, and to various degrees attempting to establish relationships between the two. Various paradigms and tasks are being employed to measure and assess many skills and abilities. Measurement instruments are being designed to be sufficiently sensitive to the behaviors of interest (e.g., microcomputer programs that employ the msec. as the unit of measurement for reaction time tasks, and GO/NO GO behavioral checklists for gross behaviors, such as replacing a wheel-bearing on a 2 1/2 ton truck).

Project A may be considered part of this concerted effort, with its concentration on identifying and furthering the conceptual and technological breakthroughs in human assessment and classification. As Hakel (1986) states, "Every major issue in the science and practice of making personnel decisions is being addressed" (p. 373).

But as Hakel also points out, "Project A won't have all the answers and it won't put the rest of us out of business." One may argue that the employment of other experimental paradigms for both the predictor and criterion sides of performance assessment by the other services suggests that Hakel is correct. There is an information-rich and burgeoning tool kit that may be called upon for reliable human assessment.

Recognition of the information that the tools (both conceptual and technological) were originally designed to tackle may provide a guide for identifying what questions each service is or is not answering. This report may provide some help toward identifying the overlaps and distinctions among the services by identifying their respective methodologies and constructs of interest, which, it may be argued, are reflecting the questions and answers on which their respective efforts focus.

REFERENCES

Annett, M. (1970). A classification of hand preferences by association analysis. British Journal of Psychology, 61, 303-321.

Bone, K. (1978). The relation between field dependency, visual search, and structure of displays. Ergonomics, 21, 383-388.

Broadbent, D. E. (1958). Perception and Communication, New York: Pergamon Press.

Christal, R. E. (1985). New cognitive tasks being evaluated by TTCP Services. Unpublished manuscript.

Christal, R. E. (1984). New cognitive tasks being evaluated by TTCP Services. Unpublished manuscript.

Cooper, L. A. (1976). Individual differences in visual comparison processes. Perception & Psychophysics, 19, 433-444.

Cooper, L. A. (1980). Spatial information processing: Strategies for research. In R. E. Snow, P. Federico, & W. E. Montague (Eds.). Aptitude, learning, and instruction: Vol 1. Cognitive process analyses of aptitude. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Dlhopolsky, J. G. (1982). Software synchronizing of video displays and z-80 processing in the Model III TRS-80. Behavior Research Methods & Instrumentation, 14, 539-544.

Eaton, N. K., Goer, M. H., Harris, J. H., & Zook, L. M. (1985). Improving the Selection, Classification and Utilization of Army Enlisted Personnel: Annual Report, 1984 Fiscal Year (Technical Report 660). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Farrell, B. (1985). "Same" - "Different" Judgments: A review of current controversies in perceptual comparisons. Psychological Bulletin, 98, 419-456.

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. Human Factors, 24, 173-183.

Gopher, D., & Kahneman, D. (1971). Individual differences in attention and the prediction of flight criteria. Perceptual and Motor Skills, 33, 1335-1342.

Hakel, M. D., (1986). Personnel Selection and Placement. Annual Review of Psychology, 37, 351-380.

Hardyk, C., & Petrinovich, L. F. (1977). Left-handedness. Psychological Bulletin, 84 385-407.

Hunt, E. (1978). Mechanics of verbal ability. Psychological Review, 85, 109-130.

Hunt, E., & Pellegrino, J. (1985). Using interactive computing to expand intelligence testing: A critique and prospective. Intelligence, 9, 207-236.

Jensen, A. R., & Munro, E. (1979). Reaction time, movement time and intelligence. Intelligence, 3, 121-126.

Kahneman, D. (1973). Attention and Effort. Englewood Cliffs, NJ: Prentice-Hall.

Kosslyn, S. M. (1981). The medium and the message in mental imagery. Psychological Review, 88, 46-66.

Krueger, L. F. (1978). A theory of perceptual matching. Psychological Review, 85, 278-304.

Lincoln, C. E., & Lane, D. M. (1980). Reaction time measurement errors resulting from the use of CRT screens. Behavior Research Methods and Instrumentation, 12, 55-57.

Logan, G. D. (1985). Skill and automaticitty: Relations, implications and future decisions. Canadian Journal of Psychology, 39, 367-386.

Long, G. M. (1972). Field dependency-independency: A review of the literature. Pensacola, Florida: Naval Aerospace Medical Research Laboratory, Naval Aerospace Medical Institute. AD-751569.

Longstreth, L. E. (1984). Jensen's reaction time investigations of intelligence: A critique: Intelligence, 8, 139-160.

McHenry, J. J., & McGue, M. K. (1975, August). Problems, issues and results in the development of computerized measures. Paper presented at the annual meeting of the American Psychological Association. Los Angeles, CA.

Mihal, W. L., & Barrett, G. V. (1976). Individual differences in perceptual information processing and their relation to automobile accident involvement. Journal of Applied Psychology, 61, 229-233.

Mos, L., Wardell, D., & Royce, J. R. (1974, January). A factor analysis of some measures of cognitive style. Multivariate Behavioral Research, 47-57.

Nickerson, R. S. (1972). Binary classification reaction time: A review of some studies of human information processing capabilities. Psychonomic Monograph Supplements, 4 (17, Whole No. 65).

Peterson, N. G. (1985, October). Mapping predictors to a criterion space: Overview. Paper presented at the annual meeting of the Military Testing Association. San Diego, CA.

Peterson, N. G. (Ed.) Development and Field Test of the Trial Battery for Project A. Technical Report ___, IN PROCESS.

Posner, M. I. (1978). Chronometric explorations of mind. Hillsdale, NJ: Lawrence Erlbaum Associates.

Posner, M., Boies, S., Eichelman, W., & Taylor, R. (1969). Retention of visual and name codes of single letters. Journal of Experimental Psychology Monographs, 79 (1, Pt. 2).

Rogosa, D., Brandt, D., & Zomowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 92, 726-748.

Rosse, R. L. (1985, August). Advantages and problems with using portable computers for personnel management. Paper presented at the annual meeting of the American Psychological Association, Los Angeles, CA.

Schneider, W. (1985, November). A quantitative model of controlled and automatic processing. Paper presented at the annual meeting of the Psychonomic Society. Boston, MA.

Sternberg, R. J. (1977). Component process in analogical reasoning. Psychological Review, 84, 353-378.

Thornton, C. L., Barrett, G. V., & Davis, J. A. (1968). Field dependence and target identification. Human Factors, 10, 493-496.

Toquam, J. L., Dunnette, M. D., Corpe, V. A., McHenry, J. J., Keys, M. A., McGue, M. K., Houston, J. S., Russell, T. L., & Hansen, M. A. (1985, August). Development of cognitive/perceptual measures: Supplementing the ASVAB. Paper presented at the annual meeting of the American Psychological Association. Los Angeles, CA.

Tucker, L. R. (1966). Learning theory and multivariate experiment: Demonstration by determination of generalized learning curves. Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally.

Zeaman, D., & Kaufman, H. (1955). Individual differences and theory in a motor learning task. Psychonomic Monographs: General and Applied, 69, (1, Whole No. 391).